hdfs fsck命令查看HDFS文件对应的文件块信息(Block)和位置信息(Locations)

在HDFS中，提供了fsck命令，用于检查HDFS上文件和目录的健康状态、获取文件的block信息和位置信息等。

fsck命令必须由HDFS超级用户来执行，普通用户无权限。

```
1.  [hadoop@dev ~]$ hdfs fsck
2.  Usage: DFSck  [-list-corruptfileblocks | [-move | -delete | -openforwrite] [-fi
    les [-blocks [-locations | -racks]]]]
3.           start checking from this path
4.         -move    move corrupted files to /lost+found
5.         -delete delete corrupted files
6.         -files  print out files being checked
7.         -openforwrite   print out files opened for write
8.         -includeSnapshots     include snapshot data if the given path indicat
    es a snapshottable directory or there are snapshottable directories under it
9.         -list-corruptfileblocks print out list of missing blocks and files they
      belong to
10.        -blocks print out block report
11.        -locations      print out locations for every block
12.        -racks  print out network topology for data-node locations
```

下面介绍每一个选项的含义及用法。

查看文件中损坏的块（-list-corruptfileblocks）

```
1.  [hadoop@dev ~]$ hdfs fsck /hivedata/warehouse/liuxiaowen.db/lxw_product_names/
    -list-corruptfileblocks
2.  The filesystem under path '/hivedata/warehouse/liuxiaowen.db/lxw_product_name
    s/' has 0 CORRUPT files
```

将损坏的文件移动至/lost+found目录（-move）

```
1.  [hadoop@dev ~]$ hdfs fsck /hivedata/warehouse/liuxiaowen.db/lxw_product_names/
    art-00168 -move
2.  FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /hivedata/wa
    ehouse/liuxiaowen.db/lxw_product_names/part-00168 at Thu Aug 13 09:36:35 CST 20
```

```
 3.    15
 3.    .Status: HEALTHY
 4.     Total size:      13497058 B
 5.     Total dirs:      0
 6.     Total files:     1
 7.     Total symlinks:              0
 8.     Total blocks (validated):       1 (avg. block size 13497058 B)
 9.     Minimally replicated blocks:    1 (100.0 %)
10.     Over-replicated blocks:         0 (0.0 %)
11.     Under-replicated blocks:        0 (0.0 %)
12.     Mis-replicated blocks:          0 (0.0 %)
13.     Default replication factor:     2
14.     Average block replication:      2.0
15.     Corrupt blocks:                 0
16.     Missing replicas:               0 (0.0 %)
17.     Number of data-nodes:           15
18.     Number of racks:                1
19.    FSCK ended at Thu Aug 13 09:36:35 CST 2015 in 1 milliseconds
20.
21.
22.    The filesystem under path '/hivedata/warehouse/liuxiaowen.db/lxw_product_names/
       part-00168' is HEALTHY
```

删除损坏的文件（-delete）

```
 1.    [hadoop@dev ~]$ hdfs fsck /hivedata/warehouse/liuxiaowen.db/lxw_product_names/p
       art-00168 -delete
 2.    FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /hivedata/war
       ehouse/liuxiaowen.db/lxw_product_names/part-00168 at Thu Aug 13 09:37:58 CST 20
       15
 3.    .Status: HEALTHY
 4.     Total size:      13497058 B
 5.     Total dirs:      0
 6.     Total files:     1
 7.     Total symlinks:              0
 8.     Total blocks (validated):       1 (avg. block size 13497058 B)
 9.     Minimally replicated blocks:    1 (100.0 %)
10.     Over-replicated blocks:         0 (0.0 %)
11.     Under-replicated blocks:        0 (0.0 %)
12.     Mis-replicated blocks:          0 (0.0 %)
13.     Default replication factor:     2
14.     Average block replication:      2.0
15.     Corrupt blocks:                 0
16.     Missing replicas:               0 (0.0 %)
17.     Number of data-nodes:           15
18.     Number of racks:                1
19.    FSCK ended at Thu Aug 13 09:37:58 CST 2015 in 1 milliseconds
20.
21.
```

```
22.  The filesystem under path '/hivedata/warehouse/liuxiaowen.db/lxw_product_names/
     part-00168' is HEALTHY
```

检查并列出所有文件状态（-files）

```
1.   [hadoop@dev ~]$ hdfs fsck /hivedata/warehouse/liuxiaowen.db/lxw_product_names/
     -files
2.   FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /hivedata/war
     ehouse/liuxiaowen.db/lxw_product_names/ at Thu Aug 13 09:39:38 CST 2015
3.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/ dir
4.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/_SUCCESS 0 bytes, 0
     block(s):  OK
5.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00000 13583807 bytes,
     1 block(s):  OK
6.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00001 13577427 bytes,
     1 block(s):  OK
7.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00002 13588601 bytes,
     1 block(s):  OK
8.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00003 13479213 bytes,
     1 block(s):  OK
9.   /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00004 13497012 bytes,
     1 block(s):  OK
10.  /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00005 13557451 bytes,
     1 block(s):  OK
11.  /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00006 13580267 bytes,
     1 block(s):  OK
12.  /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00007 13486035 bytes,
     1 block(s):  OK
13.  /hivedata/warehouse/liuxiaowen.db/lxw_product_names/part-00008 13481498 bytes,
     1 block(s):  OK
14.  ...
```

检查并打印正在被打开执行写操作的文件（-openforwrite）

```
1.   [hadoop@dev ~]$ hdfs fsck /hivedata/warehouse/liuxiaowen.db/lxw_product_names/
     -openforwrite
2.   FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /hivedata/war
     ehouse/liuxiaowen.db/lxw_product_names/ at Thu Aug 13 09:41:28 CST 2015
3.
     ..........................................................................
4.
     ..........................................................................
5.   .Status: HEALTHY
6.    Total size:    2704782548 B
7.    Total dirs:    1
8.    Total files:   201
9.    Total symlinks:            0
10.   Total blocks (validated):      200 (avg. block size 13523912 B)
```

```
11.   Minimally replicated blocks:   200 (100.0 %)
12.   Over-replicated blocks:       0 (0.0 %)
13.   Under-replicated blocks:      0 (0.0 %)
14.   Mis-replicated blocks:        0 (0.0 %)
15.   Default replication factor:   2
16.   Average block replication:    2.0
17.   Corrupt blocks:               0
18.   Missing replicas:             0 (0.0 %)
19.   Number of data-nodes:         15
20.   Number of racks:              1
21.  FSCK ended at Thu Aug 13 09:41:28 CST 2015 in 10 milliseconds
22.
23.  The filesystem under path '/hivedata/warehouse/liuxiaowen.db/lxw_product_name
     s/' is HEALTHY
```

打印文件的Block报告（-blocks）

需要和-files一起使用。

```
1.   [hadoop@dev ~]$  hdfs fsck /logs/site/2015-08-08/lxw1234.log -files -blocks
2.   FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /logs/site/20
     15-08-08/lxw1234.log at Thu Aug 13 09:45:59 CST 2015
3.   /logs/site/2015-08-08/lxw1234.log 7408754725 bytes, 56 block(s):  OK
4.   0. BP-1034052771-172.16.212.130-1405595752491:blk_1075892982_2152381 len=134217
     728 repl=2
5.   1. BP-1034052771-172.16.212.130-1405595752491:blk_1075892983_2152382 len=134217
     728 repl=2
6.   2. BP-1034052771-172.16.212.130-1405595752491:blk_1075892984_2152383 len=134217
     728 repl=2
7.   3. BP-1034052771-172.16.212.130-1405595752491:blk_1075892985_2152384 len=134217
     728 repl=2
8.   4. BP-1034052771-172.16.212.130-1405595752491:blk_1075892997_2152396 len=134217
     728 repl=2
9.   5. BP-1034052771-172.16.212.130-1405595752491:blk_1075892998_2152397 len=134217
     728 repl=2
10.  6. BP-1034052771-172.16.212.130-1405595752491:blk_1075892999_2152398 len=134217
     728 repl=2
11.  7. BP-1034052771-172.16.212.130-1405595752491:blk_1075893000_2152399 len=134217
     728 repl=2
12.  8. BP-1034052771-172.16.212.130-1405595752491:blk_1075893001_2152400 len=134217
     728 repl=2
13.  9. BP-1034052771-172.16.212.130-1405595752491:blk_1075893002_2152401 len=134217
     728 repl=2
14.  10. BP-1034052771-172.16.212.130-1405595752491:blk_1075893003_2152402 len=13421
     7728 repl=2
15.  11. BP-1034052771-172.16.212.130-1405595752491:blk_1075893004_2152403 len=13421
     7728 repl=2
16.  12. BP-1034052771-172.16.212.130-1405595752491:blk_1075893005_2152404 len=1342
     7728 repl=2
```

```
17.   13. BP-1034052771-172.16.212.130-1405595752491:blk_1075893006_2152405 len=13421
      7728 repl=2
18.   14. BP-1034052771-172.16.212.130-1405595752491:blk_1075893007_2152406 len=13421
      7728 repl=2
19.   ...
```

其中，/logs/site/2015-08-08/lxw1234.log 7408754725 bytes, 56 block(s): 表示文件的总大小和block数；

0. BP-1034052771-172.16.212.130-1405595752491:blk_1075892982_2152381 len=134217728 repl=2

1. BP-1034052771-172.16.212.130-1405595752491:blk_1075892983_2152382 len=134217728 repl=2

2. BP-1034052771-172.16.212.130-1405595752491:blk_1075892984_2152383 len=134217728 repl=2

前面的0. 1. 2.代表该文件的block索引，56的文件块，就从0-55;

BP-1034052771-172.16.212.130-1405595752491:blk_1075892982_2152381表示block id；

len=134217728 表示该文件块大小；

repl=2 表示该文件块副本数；

打印文件块的位置信息（-locations）

需要和-files -blocks一起使用。

```
1.    [hadoop@dev ~]$  hdfs fsck /logs/site/2015-08-08/lxw1234.log -files -blocks -lo
      cations
2.    FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /logs/site/20
      15-08-08/lxw1234.log at Thu Aug 13 09:45:59 CST 2015
3.    /logs/site/2015-08-08/lxw1234.log 7408754725 bytes, 56 block(s):  OK
4.    0. BP-1034052771-172.16.212.130-1405595752491:blk_1075892982_2152381 len=134217
      728 repl=2 [172.16.212.139:50010, 172.16.212.135:50010]
5.    1. BP-1034052771-172.16.212.130-1405595752491:blk_1075892983_2152382 len=134217
      728 repl=2 [172.16.212.140:50010, 172.16.212.133:50010]
6.    2. BP-1034052771-172.16.212.130-1405595752491:blk_1075892984_2152383 len=134217
      728 repl=2 [172.16.212.136:50010, 172.16.212.141:50010]
7.    3. BP-1034052771-172.16.212.130-1405595752491:blk_1075892985_2152384 len=134217
      728 repl=2 [172.16.212.133:50010, 172.16.212.135:50010]
8.    4. BP-1034052771-172.16.212.130-1405595752491:blk_1075892997_2152396 len=134217
      728 repl=2 [172.16.212.142:50010, 172.16.212.139:50010]
9.    5. BP-1034052771-172.16.212.130-1405595752491:blk_1075892998_2152397 len=134217
      728 repl=2 [172.16.212.133:50010, 172.16.212.139:50010]
10.   6. BP-1034052771-172.16.212.130-1405595752491:blk_1075892999_2152398 len=134217
      728 repl=2 [172.16.212.141:50010, 172.16.212.135:50010]
```

```
11.   7. BP-1034052771-172.16.212.130-1405595752491:blk_1075893000_2152399 len=134217
      728 repl=2 [172.16.212.144:50010, 172.16.212.142:50010]
12.   8. BP-1034052771-172.16.212.130-1405595752491:blk_1075893001_2152400 len=134217
      728 repl=2 [172.16.212.133:50010, 172.16.212.138:50010]
13.   9. BP-1034052771-172.16.212.130-1405595752491:blk_1075893002_2152401 len=134217
      728 repl=2 [172.16.212.140:50010, 172.16.212.134:50010]
14.   ...
```

和打印出的文件块信息相比，多了一个文件块的位置信息：[172.16.212.139:50010, 172.16.212.135:50010]

打印文件块位置所在的机架信息（-racks）

```
1.    [hadoop@dev ~]$  hdfs fsck /logs/site/2015-08-08/lxw1234.log -files -blocks -lo
      cations -racks
2.    FSCK started by hadoop (auth:SIMPLE) from /172.16.212.17 for path /logs/site/20
      15-08-08/lxw1234.log at Thu Aug 13 09:45:59 CST 2015
3.    /logs/site/2015-08-08/lxw1234.log 7408754725 bytes, 56 block(s):  OK
4.    0. BP-1034052771-172.16.212.130-1405595752491:blk_1075892982_2152381 len=134217
      728 repl=2 [/default-rack/172.16.212.139:50010, /default-rack/172.16.212.135:50
      010]
5.    1. BP-1034052771-172.16.212.130-1405595752491:blk_1075892983_2152382 len=134217
      728 repl=2 [/default-rack/172.16.212.140:50010, /default-rack/172.16.212.133:50
      010]
6.    2. BP-1034052771-172.16.212.130-1405595752491:blk_1075892984_2152383 len=134217
      728 repl=2 [/default-rack/172.16.212.136:50010, /default-rack/172.16.212.141:50
      010]
7.    3. BP-1034052771-172.16.212.130-1405595752491:blk_1075892985_2152384 len=134217
      728 repl=2 [/default-rack/172.16.212.133:50010, /default-rack/172.16.212.135:50
      010]
8.    4. BP-1034052771-172.16.212.130-1405595752491:blk_1075892997_2152396 len=134217
      728 repl=2 [/default-rack/172.16.212.142:50010, /default-rack/172.16.212.139:50
      010]
9.    5. BP-1034052771-172.16.212.130-1405595752491:blk_1075892998_2152397 len=134217
      728 repl=2 [/default-rack/172.16.212.133:50010, /default-rack/172.16.212.139:50
      010]
10.   ...
```

和前面打印出的信息相比，多了机架信息：[/default-rack/172.16.212.139:50010, /default-rack/172.16.212.135:50010]

转载请注明：lxw的大数据田地 » hdfs fsck命令查看HDFS文件对应的文件块信息(Block)和位置息(Locations)